Martijn Catrycke, Giulio Antonio Abbo, Maria Jose Pinto-Bernal and Tony Belpaeme

Multi-party Open-Ended Conversations

with the Furhat Robot

IDLab-AIRO
Ghent University – imec
Belgium

Implementation
Modular architecture
Event based

Evaluation
30 subjects
Group conversation
Parallel conversations

Conversation Manager
☞ Dialogue generation
☞ Addressee injection

Turn Taking
☞ Decide when to talk

Face Tracking & Voice Recognition
☞ Speaker position
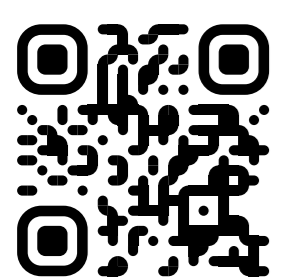☞ Absolute speaker recognition

Interactions Manager
Text to speech ☞
Interrupt handling ☞
Attend to user ☞

Diarisation
Speech to text ☞
Track last spoken ☞
sentence

Speaker Awareness
Voice direction of arrival ☞
Turn changes detection ☞

Left: link to the paper. **Below**: frequencies of the system's latencies. **Bottom**: voice recognition, face recognition and addressee injection performances. **Right**: how many times each topic was mentioned during quantitative evaluations, with positive and negative connotations, in the two settings.

Voice Recognition
incorrect 2.8%
correct 21.9%
null 75.3%

Face Recognition
correct 85.9%
null 0.8%
incorrect 13.4%

Addressee Injection
correct 85.2%
null 4.1%
incorrect 10.8%